DOCUMENT RESUME

ED 314 749 CS 212 188

AUTHOR Anderson, Philip M.
TITLE Evaluative Response to Poetic Convention at Four Grade Levels.
PUB DATE 90
NOTE 25p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Adolescents; Evaluation Research; Grade 5; Grade 7; Grade 9; Grade 11; Intermediate Grades; Literary Devices; Literature Appreciation; *Poetry; *Reader Response; Reading Attitudes; Secondary Education; Student Reaction; Text Structure
IDENTIFIERS New York City Board of Education; *Reader Preferences; Text Factors

ABSTRACT

A study examined the poetic conventions of rhyme, meter, and diction to determine whether central tendencies and differences in evaluative response at various stages of late childhood and adolescence were discernible. Data were gathered from students of four grade levels (5, 7, 9, 11) enrolled in New York City public schools (359 students in all). Students (with varied knowledge of poetry) were asked to give best/worst evaluative responses to four versions each of three nonnarrative poems chosen from contemporary school anthologies. The versions were the original and three altered versions which, respectively, unrhymed the poem, disturbed the meter, and replaced poetic diction with prosaic diction. Results revealed three patterns of central tendency in the responses. First, a central tendency in evaluative response to the individual poems at each grade level was evident. In each case more than 33% of the respondents chose one altered version over the other three for "best" and for "worst." Second, responses indicated a central tendency in evaluation of altered poetic conventions across the poems. The combined percentages of original and altered diction "best" choices was larger than the combined percentages for rhyme and meter as "best," while the obverse was true for the "worst" versions. Third, a central tendency was evident in individual preferences regarding preferred conventions or forms. More than 21% of students at each of the grade levels selected the same matrix more than once. No clear-cut developmental schema was discernible from the data. (Four tables of data are included; 17 references are attached.) (KEH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Philip M. Anderson
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED314749

Evaluative Response to Poetic Convention at Four Grade Levels

Philip M. Anderson

Associate Professor, Department of Secondary Education and Youth Services

Queens College of the City University of New York

Flushing, NY 11367-0904

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Two main methodological trends characterize research on response to literature. One is the protocol analysis method popularized by I. A. Richards (1929), which has guided most of the research in the field (Squire, 1964; Holland, 1975; Petrosky, 1976; Bleich, 1978; Applebee, 1978; Beach and Wendler, 1987). Generally, participants in protocol analysis research write or speak a response to a literary text, and the response is analyzed to infer various characteristics of the responders or the response. The theoretical grounding of protocol analysis in the above studies has been either psychoanalytic, hermeneutic, or psycholinguistic, based on theorists ranging from Freud to James Britton. And, though Richards' study analyzed response to poetry, most of these studies have described and analyzed response to prose.

A second stream of studies, initiated by Allan Abbott and M. R. Trabue (1921), seeking to evaluate response before it is articulated verbally, has guided a smaller number of researchers (Hansson, 1964; Purves, 1973, 1981). Investigating what some researchers refer to as "passive ability," unarticulated response research seeks to avoid the interpretation and translation problems of protocol analysis, "to try to go beyond such descriptions and statements and try to get closer to the response itself" (Hansson, 1985, pp. 217, 213). Unarticulated response research employs various methodologies: Abbott and Trabue (1921) asked subjects to choose the best and worst among four altered versions of the same poem, one the original text; Hansson (1964) charted responses on a semantic differential scale; and Purves (1973) asked readers to choose among critical questions they would prefer to answer about a text.

CS212188

BEST COPY AVAILABLE

The Response to Literature report from the Dartmouth conference (Squire, 1968) provided theoretical support for the unarticulated response approach:

> "A sense of literary form . . . may become articulate finding expression in comment and criticism, but equally it may not; and this, as pedants, we find very difficult to admit" (Britton, p. 5); [and,] ". . . overt response (verbal, etc.) may indicate very little of the inner response" (Harding, p. 11).

Regardless of the theoretical influence of Dartmouth, Applebee (1977) demonstrated (in a review of literary response studies since 1968) that protocol analysis was the predominant choice of method for response research for the following decade.

A common feature of unarticulated response research is a methodological focus on textual variables. Readers respond to textual variations or respond within given semantic differentials; researchers chart the group's responses to ascertain whether there are patterns of response to a text. The approach aims at assessing normative response to literary texts, and discovering whether textual variables affect response. The most important finding of unarticulated response research is the idea of a "central tendency" in response to a particular text:

> Using the semantic differential with a variety of readers from all walks of life, [Hansson (1964)] showed that there was a central tendency in the response to a poem. Only when the readers began to articulate their responses did diversity enter, for then readers began to focus on different aspects of the poem, began to make critical statements (Purves, 1979, p. 809).

Purves' IEA study (1973) found "that each text is unique not simply in its meaning and emotion but in the critical perspective that it elicited from students, and this across a broad spectrum of cultures and educational systems" (Purves, 1979, p. 810). Abbott and Trabue's study (1921) found a similar central tendency with regard to student evaluative judgment in response to altered versions of poems, but only with the upper grades of the high school and in college. With younger respondents, down to grade five, the response appeared idiosyncratic.

Background of the Study

The present study is a version of Abbott and Trabue's (1921) study in which students in grades five through graduate school were asked to choose the best and worst from four different versions of twenty-six poems re-written by the experimenters. Abbott and Trabue's criteria for altering the texts were what they deemed the "essential characteristics" of poems: diction, meter, and tone. The altered diction version made the poem more prosaic in diction, the altered meter version disturbed the rhythm of the poem; the altered tone version made the poem more "sentimental." Students were asked to choose the best and worst versions from among the original and the three altered versions of each poem. Abbott and Trabue found several central tendencies: a common dislike of the altered meter version at almost all grade levels, a general approval of the "sentimental" version in the high school and college, and a consistent identification of the original poem as the best version only in college and graduate school. They are inconclusive about the ability of students in grades 5 - 11 to evaluate poetry in general, but suggest that either the instrument or the students' judgments are unreliable.

There are several methodological problems with Abbott and Trabue's study that complicate their findings. First, participants were asked to respond to a large number of poems at one sitting; the poems were divided into two groups of thirteen (actually fifty-two, four versions of each) poems each and responses were recorded during two forty-minute sessions. As well, the original poems in the study were of different lengths and types. Secondly, the experimenters substantially altered each version when re-writing the poems, thereby confounding textual variables. Each version changed enough that it dramatically altered other text variables (i.e., more than just the meter had been altered in the altered meter poems).

And finally, the criteria of "essential characteristics" of poetry did not produce sufficiently parallel text variables. The "tone" alteration, making the poem more "sentimental," is especially difficult to measure, and involved a different type of alteration than meter and diction changes. An earlier reviewer of Abbott and Trabue also suggested that the versions of the poems are "too radically altered" to be comparable measures (Eppel, 1950, p. 111). The present study

chose "poetic conventions" (meter, rhyme, and diction) as the criteria for textual alteration. Altering the meter involved adding a syllable or deleting a syllable sufficient to make the meter uneven (e.g., "We never know how high we are" became "We never know how high up we are"). Altered rhyme involved replacing rhyming words at the end of a line with non-rhyming synonyms. The altered diction version involved replacing some of the more conventionally poetic usages in the poems with more prosaic synonyms (e.g., "shan't" becomes "won't"). The number of changes in each of the versions numbered no more than eight words.

Rather than make each version of the poem quite different from the others, as Abbott and Trabue did, the present study attempted to make the versions as similar as possible. Not only does the similarity of the versions of the poems in the present study call for more discriminating judgment by the respondent, but limiting the degree of change in the rewrites minimizes the inevitable additional alterations to the stated textual variable that occur when the poem is re-written. In the present study, students at all grade levels responded to the same three complete non-narrative poems, one of six lines and two of eight lines. Abbott and Trabue's decision to use different types of poems and partial poems seems to imply they thought the "essential characteristics" of poems were not dependent on the structure of the poem.

Students at grade levels 5, 7, 9, and 11 were chosen because they represent the normal range of students taught by middle and high school English teachers. As well, this age range represents the group for which Abbott and Trabue's study was inconclusive. Both Abbott and Trabue's study and this study chose poems from contemporary school anthologies.

Purpose of the Study

This study is an attempt to ascertain whether the poetic conventions of rhyme, meter, and diction are factors affecting unarticulated evaluative response to non-narrative poetry. Response patterns of <u>best</u> and <u>worst</u> choices at each of four grade levels (5, 7, 9, 11) will reveal whether there are any central tendencies in evaluative response, and whether there are differences in evaluative response at various stages of late childhood and adolescence.

Method

Participants

The participants are students in New York City public schools, four classes at each of four grade levels: grades five (N = 98), seven (N = 94), nine (N = 85), and eleven (N = 82). Two classes of fifth graders attended a middle school (two different teachers, Class 1: N = 25; Class 2: N = 28) and two classes attended an elementary school (two different teachers, Class 3: N = 22, Class 4: N = 24). The seventh graders attended two different middle schools (one teacher at each school, Classes 1, 2: N = 27, 25 [classes listed together signify the same teacher for both classes]; Classes 3, 4: N = 24, 18). The ninth graders attended two different high schools (one teacher in each school, Classes 1,2: N = 15, 24; Classes 3,4: 26, 23). The eleventh graders attended one high school (three different teachers, Class 1: N = 14; Classes 2,3: N = 26, 25; Class 4: N = 19).

Teachers of the classes were asked to describe their students' prior experiences with poetry, types of poems the teachers taught during the year, and their methods of teaching poetry. Questions concerning their students also included perceived ability level characterizations.

In general, grade five students (representing a variety of ethnic backgrounds, as did all the school populations in the study) were characterized as average or above average in reading ability and most were designated as college-bound. All students were characterized as having prior knowledge of poetry. The four teachers cited varying goals for teaching poetry and, with the exception of Class 1, provided only sporadic experiences, generally given as homework. Lyric and narrative poetry were named as the types of poetry studied.

Grade seven classes split between a selective arts and humanities-based middle school and a regular middle school. In the selective classes (Classes 1 and 2, taught by the same teacher), the students engaged in intensive study of poetry during the school year. Students in the other two classes were described by their teacher as college-bound students with some poetry background.

Classes 1 and 2 of the grade nine classes were characterized as having no prior knowledge of poetry, and all four classes were characterized as receiving little instruction in poetry, and most of that as homework assignments.

In grade eleven, Class 1 was described as having little prior knowledge of poetry, and the teacher taught only a brief poetry unit at the end of the spring term. Classes 2 and 3 (both taught by the same teacher) received intensive instruction in poetry from Chaucer to Hardy during an historical survey of literature. Class 4's teacher also provided intensive teaching of poetry, including the sonnet and ballad forms.

None of the respondents, at any grade level, had studied any of the three poems in this study during the year, though "The Pasture" was in the anthology of the selective middle school (Grade 7, Classes 1 and 2).

Instrument

The instrument consisted of four versions (three altered versions and the original, numbered 1 - 4) of Robert Frost's "The Pasture (A)," Alfred Tennyson's "The Eagle (B)," and Emily Dickinson's "We Never Know How High (C)" identified by title and the letter A, B, or C. The altered versions 1) unrhymed the poem, 2) disturbed the meter, and 3) replaced poetic diction with prosaic diction. In altering the poems, the criterion was to change the poem as little as possible from the original, while consistently altering each textual variable.

All four versions of each poem were typed on a single page, each sequence in a different order ("Pasture" order: altered rhyme, altered diction, original, altered meter; "Eagle": R, O, M, D; "We Never": M, O, R, D). The three poems were stapled into a packet; the poems were rotated in the packets (A, B, C; B, C, A; C, A, B), with one-third of each class responding to each sequence, to account for any fatigue factor in response. Participants chose a best and a worst version for each set of versions of the three poems because critical judgment involves both positive and negative choices.

Procedure

A separate answer sheet (the "[A]" version is shown below, for the packet that placed the "A" poem, "The Pasture," first), corresponding to the order of the poems in that respondent's packet of poems, provided the following instructions, read aloud by the experimenter:

> Answer the following questions about the sets of poems you will read by placing the number of the version you choose from each set on the lines below. We will read and evaluate one set of poems at a time. Do the set of poems on the top page first. Don't do the second and third sets until you are told to do so.
>
> 1) Which version of the poem is the best poem? It may be the one you like best or it may not be the one you like best. Which is the best poem?
>
> 2) Which version of the poem is the worst poem? It may be the one you like the least or it may not be. Which is the worst poem?
>
> Put the number of the version you choose in the blanks below. Select only one "best" and one "worst" version for each set of poems. Make sure you are referring to the correct poem (A, B, or C; look at the top of the page on which the poems appear). Do not write on the poems themselves.
>
> Poem A: "The Pasture"
>
> 1) Best Version # _____ 2) Worst Version # _____
>
> Poem B: "The Eagle"
>
> 1) Best Version # _____ 2) Worst Version # _____
>
> Poem C: "We Never Know How High"
>
> 1) Best Version # _____ 2) Worst Version # _____

The answer blanks for the other two packets (B, C, A; C, A, B) are in matching sequences. The procedure was completed in approximately twenty minutes in each class (though one seventh

grade class completed the task in only ten minutes). Responses were collected over a period of four weeks at the end of the school year.

Results

The data is organized around grade level responses to each poem, and the accompanying tables (matrix of best / worst choices) show both the best and worst evaluative responses to each poem at the four grade levels (figs. 1-3). Chi square values for each of the matrices are significant at the $p \leq .001$ level ($df = 9$).

Theoretically, only responses to the same poem can be compared, since the focus of the study is textual variables and each poem represents a different response situation. However, similar grade level responses to poetic convention across the poems (see Table 4), and repeated choices of a particular matrix of poetic conventions by individual respondents, are also reported to account for any patterns of learned group or personal responses to particular conventions.

In tabulating the results, responses in which a student did not indicate both a best and worst choice were not counted for the particular poem. As a result, there are slightly different totals for each grade level across the poems.

"The Eagle"

Responses to the versions of "The Eagle" (see Table 1) show the altered diction (D) version a clear best choice in grades 5, 7, and 9 (39/97 = 40.21%; 32/94 = 34.04%; 53/88 = 60.23%), while in the 11th grade the original (O) and altered diction (D) were almost equal choices (O: 30/82 = 36.59%; D: 29/82 = 35.37%). The most frequent second choices for best version were: 5th grade, altered meter (M), 7th grade, altered rhyme (R), 9th grade, the original (O).

Altered rhyme (R) was designated worst most frequently at all grade levels (Grade 5: 56/97 = 57.73%; Grade 7: 34/94 = 36.17%; Grade 9: 55/88 = 62.5%; Grade 11: 34/82 = 41.46%). The rest of the worst choices were spread rather evenly across the other options, but

the original version attracted the smallest number of worst choices at all grade levels (though it tied with the other two choices in Grade 7).

The most frequent matrix choice (combination of best and worst choices) at all grade levels was "altered diction best/altered rhyme worst." At the 11th grade level, the "original best/altered diction worst" was the second most frequent matrix choice. Class totals were similar at all grade levels.

Insert Table 1 here

"We Never Know How High"

The response pattern for "We Never Know How High" (see Table 2) is similar to that for "The Eagle." The altered diction (D) version was the most frequent choice for best version in grades 5 (39/98 = 39.8%), 7 (31/93 = 33.33%), and 9 (39/87 = 44.83%), until it is overtaken by the original (O) in grade 11 (30/80 = 37.5%) (though the altered diction version is still an important choice within the group [25/80 = 31.25%]).

The original version was the second most frequent choice of the 5th and 9th graders, but a large percentage of the 7th graders (27/93 = 29.03%) chose the altered rhyme version as best. In grades 5, 7, and 9, the altered meter version was chosen the fewest times as best, and in the 11th grade the altered rhyme version was selected least as best.

Altered rhyme and meter versions were both frequent worst choices in grades 5 and 7 (5: R, 33/98 = 33.67%, M, 29/98 = 29.59%; 7: R, 28/93 = 30.11%, M, 33/93 = 35.48%). In grades 9 and 11, the altered rhyme version was the most frequently chosen worst choice (9: 36/87 = 41.38%; 11: 44/80 = 55.0%). The least frequently chosen worst versions were the original and the altered diction (though in grade 9 the original was chosen only one fewer time than the altered meter version).

The most frequent matrix choice in grades 5 and 9 was the "altered diction best/altered rhyme worst." In grade 7, the most frequent matrix choice was "altered diction best/altered meter worst," while in grade 11 the "altered diction best/altered rhyme worst" and "original best/altered rhyme worst" were almost equal choices (17 and 18 respectively).

Differences based on gender appeared in the grade 7 and 11 responses to "We Never Know How High." In grade 7, a much higher percentage of males (14/39 = 35.9%) than females (7/54 = 12.96%) chose the original version as best, though males and females had the same percentages for altered diction as best category (male: 13/39 = 33.33%; female: 18/54 = 33.33%). The large number of choices for the altered rhyme version as best in grade 7 was strictly a gender difference (female: 19/54 = 35.19%; male: 8/39 = 20.51%). The 7th grade patterns are not skewed by any particular class; both the select middle school students and the regular junior high students presented similar patterns of response.

In grade 11, the male responses for best were almost equally divided across the categories (O = 4, R = 6, M = 7, D = 5), while female responses showed a preference for the altered diction and original versions as best (O = 26, R = 5, M = 7, D = 20). Though there were gender differences in choices for worst, both males (11/22 = 50%) and females (33/58 = 56.9%) chose the altered rhyme version as a clear choice for worst.

Within the grade 5 and 9 classes, there was little variation in the pattern of choices by gender. Also, gender differences do not appear in responses to "The Pasture" or "The Eagle."

Insert Table 2 here

"The Pasture"

The response pattern for "The Pasture" (see Table 3) is somewhat different from the patterns for the other poems. In grade 5, the altered meter version was judged best (35/98 = 35.71%), while in grade 7, the original version was judged best (35/93 = 37.63%). The best

response patterns for grades 9 and 11 remained similar to the earlier poems: the 9th graders preferred the altered diction version (29/85 = 34.12%) and the 11th graders chose the altered diction and original versions almost equally (O: 28/80 = 35.0%, D: 30/80 = 37.5%). In the 9th grade the altered meter version received the second highest number of best choices (26/85 = 30.59%), more than the original as best choices (23/85 = 27.06%).

As for the frequency of worst choices, grade 5 selected altered rhyme most frequently (39/98 = 39.8%), while grade 7 almost equally chose the altered rhyme (32/93 = 34.41%) and altered meter (28/93 = 30.11%) versions as worst choices. The original version was chosen least as the worst version in grades 5, 7, and 9 (though in grade seven altered meter had only one more tally), and tied with the altered diction version in grade 11.

No common matrix patterns appeared in the responses to "The Pasture." But the most frequently selected matrices were in the altered rhyme and altered meter as worst columns in grades 5, 7, and 9, while in grade 11 the highest choice matrices were both in the altered meter as worst column. Also, in grades 7 and 9 a number of responses fell in the "altered diction best/original version worst" matrix (Grade 7: 10/93; Grade 9: 10/85).

Accounting for the choice of the altered meter version as the preferred best choice in grade 5 (although all four grade 5 classes rated the altered meter version either first or second as their best choice), Class 2 chose the altered meter version as best by a higher percentage (12/27 = 44.44%) than the other classes (7/25 = 28.0%, 8/22 = 36.36%, 8/24 = 33.33%). Class 2 also chose the original version as best in much higher numbers (10/27 = 37.04%) than the other classes (6/25 = 24.0%, 6/22 = 27.27%, and 4/24 = 16.67%), and ignored the altered diction version (2/27 = 7.41%) the others emphasized (10/25 = 40.0%, 7/22 = 31.82%, and 8/24 = 33.33%). Equal numbers of males and females chose the altered meter and original versions as best choices in that class. In this class as well, larger numbers of students chose the altered diction version as worst (10/27) than in other classes (4/25, 3/22, 6/24).

Insert Table 3 here

Group Response Totals

Theoretically, evaluative choices across the poems cannot be compared, but in this case response to poetic convention appears to transcend the individual texts. The tally of choices shows a central tendency toward the original and altered diction version as best choices and altered rhyme and altered meter as worst at all grade levels, except grade 5 which chose the altered meter version one more time than the original as the second most frequent choice (see Table 4). In grades 5, 7 and 9 the altered diction version is chosen best most often and the original is chosen second most often in grades 7 and 9, while in grade 11 the situation is reversed. The altered rhyme version is chosen most often as worst at all grade levels and the altered meter version is the next most frequent choice for worst at all grade levels.

Insert Table 4 here

Individual Patterns of Response

There are also cases of individuals respondents who placed two out of three combined responses in the same matrix. That is, of the three combinations of best/worst choices the individual respondent made, the combination of choices for two of the three poems were identical (respondents could not do this by simply marking the same numbers on the answer sheet, because the altered versions were in a different order for each poem). In six cases, the respondent matched on all three response matrices.

In grade 5, the number, by class group, of those who placed two out of three best/worst choices at the same place on the matrix are:

Class 1: 7/25 = 28.0% (male: 4/16, female: 3/9)

Class 2: 9/28 = 32.14% (M: 5/14, F: 4/14) (one male placed all three of his choices in the same matrix)

Class 3: 10/22 = 45.45% (M: 7/16, F: 3/6)

Class 4: 7/24 = 29.17% (M: 3/9, F: 4/15)

The total of matched responses was 33 out of 98 respondents (33.67%).

There was no pattern of combination of poems in the matched responses, but a majority of the matched responses were in the "original best/altered rhyme worst" or "altered diction best/altered rhyme worst" matrices. The third choice of the matched responses in grade five was not predictable. The one respondent who chose the same matrix of choices in all three cases chose "original best/altered diction worst."

In grade 7, the numbers of students, by class, placing two of three best/worst choices in the same place on the grid are:

Class 1: 6/27 = 22.22% (M: 0/10, F: 6/17) (one female matched on all three choices)

Class 2: 6/25 = 24.0% (M: 4/13, F: 2/12) (one male placed all three choices in the same matrix)

Class 3: 4/24 = 16.67% (M: 0/9, F: 4/15)

Class 4: 4/18 = 22.22% (M: 1/7, F: 3/10)

The total of the matched responses was 20 out of 94 respondents (21.28%).

There are no particular patterns as to the combination of poems the individual students matched, but the larger part of the matched matrices are at "original best/altered rhyme worst" and "altered diction best/altered rhyme worst."

The four classes of 9th graders present a less regular distribution. Two groups tallied low numbers of matched scorers, and two groups compare favorably with the 5th and 7th grade groups:

Class 1: 2/15 = 13.33% (M: 1/6, F: 1/9),

Class 2: 8/24 = 33.33% (M: 6/15, F: 2/9),

Class 3: 9/26 = 34.62% (M: 4/11, F: 5/15) (one male matched all three)

Class 4: 1/23 = 4.35% (M: 0/7, F: 1/16).

The total of matched responses was 20 out of 88 respondents (22.73%).

The 2/15 and the 8/24 groups (Classes 1, 2) had the same teacher, and the 9/26 and 1/23 groups (Classes 3,4) were taught by the same teacher. As for matrix patterns, almost all the matched responses (including the three-way match) occurred in the "altered diction best/altered rhyme worst" matrix.

At the 11th grade level, there is a larger percentage of respondents with matched matrix choices for two poems than at the 7th and 9th grade levels:

Class 1: 8/12 = 66.67% (M: 4/6, F: 4/6)

Class 2: 5/25 = 20.0% (M: 1/8, F: 4/17, two of whom had a three way match)

Class 3: 8/25 = 32.0% (M: 1/7, F: 7/18)

Class 4: 5/18 = 27.78% (M: 0/1, F: 5/17).

The total for the eleventh grade was 26, or 31.71% of the 82 participants.

Twenty-three matched responses were in the original or altered diction best columns and the altered meter and rhyme worst columns (the other three included one on the "altered meter best/original worst" matrix, one on the "R/D" matrix and one on the "M/D" matrix)). With rare exceptions, the third choice of the matched pair respondents designated the original or the altered diction version as the best version, regardless of worst choice, even in cases where the matched pair did not choose the original or altered diction as best.

Discussion

The data reveal three patterns of central tendency in the responses. First, a central tendency in evaluative response to the individual poems at all grade levels. In each case more than 33% of the respondents chose one altered version over the other three for best and for worst. Second, responses indicated a central tendency in evaluation of altered poetic conventions across the poems. The combined percentages of original and altered diction best choices was larger than the combined percentages for rhyme and meter as best, while the obverse was true for the worst

versions. And, third, a central tendency is evident in individual preferences regarding preferred conventions or forms. More than 21% of students at each of the grade levels selected the same matrix more than once.

But, the data also reveal some patterns of dispersion within the central tendencies. The pattern of choices for "The Pasture" was different than the pattern for the other two poems in grades 5 , 7, and 9; altered meter was the most frequent best choice in grade 5, the original was the most frequent best choice in grade 7, altered diction was tied with altered rhyme as the most frequent worst choice in grade 9, and altered meter was selected as the most frequent worst choice.

With the exception of the grade seven choices for "The Pasture," only the grade eleven students chose the original as best more often than the altered diction in response to every poem, and grade seven responders picked the altered rhyme version as the second most frequent best choice in two out of three poems.

Patterns of response indicating the altered diction version as the most frequent best choice by students in grades 5, 7, and 9 suggest that vocabulary may be a factor in evaluative response for younger adolescents, at least in the case of these poems. Eleventh grade responses, along the lines of a predicted adult normative response, suggest that poetic diction does not present as large a vocabulary problem in that grade, though the altered diction as best choice rivalled the original at that grade level as well. The tendency toward choosing the altered diction version as best may indicate that the evaluative response relies on comprehension of the poem, rather than merely response to form. On the other hand, since students gave each version only a single reading without discussion, the choice of the prosaic language may be less a matter of comprehension and more a preference in style.

Patterns of response showing a clear dislike of unrhymed poetry and poetry without regular meter at all grade levels may indicate a lack of exposure to free and blank verse, or a preference for rhymed and regular verse. Also, since the poems were read silently, these choices may indicate some sub-vocalization may be taking place in reading poems.

Gender differences at two grade levels (7 and 11) in the response to the poem written by the female poet suggest further possibilities for unarticulated response research, but the present study did not consider the gender of the poet as a variable in the original design.

The pattern of matched matrix choices in grade 7 suggest that the teacher variable was less important that individual preference in those classes: Classes 1 and 2 (the 2/15 and the 8/24 groups) were both taught by one teacher, while both Classes 3 and 4 (the 9/26 and 1/23 groups) were taught by another. Though matched matrix choices tended toward the generally preferred "altered diction best/altered rhyme worst" matrix, the matched matrices were sufficiently individual in nature to suggest that personal preference was stronger than individual teacher influence. And, not all matched matrix choices showed "good" judgment, because not all the matched matrix responses to the poems fell within the majority categories. So even though some of the matched matrix choices might be accounted for by the consensus choices of the group, many did not reflect the consensus, indicating that individual preference is probably the determining factor.

On the other hand, as to school and teacher effects in general, the responses still need to be seen as "schooled" responses, and the movement toward the original as best as the students grow older, and the altered rhyme and meter as worst at all levels, is probably a function of aesthetic and linguistic experience received in school.

But, there is no clear-cut developmental schema discernible from the data, even though eleventh graders did chose the original as best for all three poems. Some other patterns based on total responses at grade level do appear. For example, percentages of worst choices at all grade levels were highest for altered rhyme, with similar percentages at all grade levels (Grade 5: 43.69%; Grade 7: 33.57%; Grade 9: 44.23%, Grade 11: 42.15%), so there appears to be an early evaluation made on the issue of unrhymed poetry by students that does not change with time.

In the case of best choices, the fifth graders chose altered meter as the second most frequent choice (though the choices in response to "The Pasture" account for the difference), while in grades 7 and 9 the students chose the original as best second to the altered diction version.

Conversely, the percentage of grade five responses for altered diction as best (35.84%) was higher than the percentage for grade seven (31.07 %), as was the combined percentage of altered diction and original choices in grade five (60.07%) higher than the combined percentage in grade seven (57.86%). The grade nine and eleven combined percentages for original and altered diction as best were higher than the lower grades but very similar to one another (Grade 9: 69.62%; Grade 11: 71.07%), though the percentage choosing the original as best was much higher in grade eleven (Grade 9: 23.08%; Grade 11: 37.19%).

One other interesting pattern in the worst choices is the similarity in the percentages of original as worst choices in grades five and eleven (Grade 5: 12.63%; Grade 11: 12.4%), both of which are lower than the percentages in grades seven and eleven (Grade 7: 17.86%; Grade 9: 18.08%). Since evaluation involves negative as well as positive judgments, the fifth graders rejection of the original as the worst choice in the same percentages as the eleventh graders may indicate a level of aesthetic judgment that does not reappear until grade eleven. Another provocative parallel between the fifth graders and the eleventh graders is the higher percentage of matched responses at those grade levels (Grade 5: 33.67%; Grade 11: 31.71%) than at the other two levels (Grade 7: 21.28%; Grade 9: 22.73%). Certain developmental characteristics of fifth graders may make them more consistent in their evaluative response, and those characteristics (or similar characteristics) may not reappear until late adolescence.

The study raises questions about the validity of theories promoting primacy of the text, the form, or the individual for aesthetic response, since all three variables appear to influence the evaluation of the poetic transaction by adolescents. The study's findings appear to reinforce Purves' assertions about the importance of the "twin ideas of central tendency and dispersion . . . to a theory of response and research in response" (1979, p. 809). As discovered by Hansson and Purves in other studies, initial responses to literature (in this case poetry) are generally consistent regardless of age, experience, or reading ability. On the other hand, the large percentages of respondents' matched matrix choices, even in the cases where the respondent did not choose the consensus response, indicate that subjective factors enter into response.

And finally, certain conventions of poetry appear to affect aesthetic judgment, especially in negative evaluations, but also in positive evaluations. Since these preferences are characteristic of the students at various levels regardless of exposure to poetry in school, they may be developmental in nature. Recognizing that students prefer more prosaic diction, while rejecting unrhymed poems, teachers may reflect those preferences in selection of poems and in teaching emphasis at various grade levels.

References

Abbott, A. A., & Trabue, M. R. (1921). A measure of the ability to judge poetry. Teachers College Record 22, 101-126.

Applebee, A. N. (1977). ERIC / RCS report: the elements of response to a literary work: what we have learned. Research in the Teaching of English 11, 255-271.

Applebee, A. N. (1978). The child's concept of story. Chicago: University of Chicago Press.

Beach, R. and Wendler, L. (1987). Developmental differences in response to a story. Research in the Teaching of English 21, 286-297.

Bleich, D. (1978). Subjective criticism. Bloomington: Indiana University Press.

Eppel, E. M. (1950). A new test of poetry discrimination. The British Journal of Educational Psychology 20, 111-116.

Hansson, G. (1964). Dikt i profil (Poetry in profile). Göteborg: Akademiförlaget / Gumperts.

Hansson, G. (1973). Some types of research on response to literature. Research in the Teaching of English 7, 260-284.

Hansson, G. (1985). Verbal scales in research on response to literature. In C. Cooper (Ed.), Researching response to literature and the teaching of literature: points of departure (pp. 212-232). Norwood, NJ: Ablex.

Holland, N. (1975). Five readers reading. New Haven: Yale University Press.

Petrosky, A. (1976). The effects of reality perception and fantasy on response to literature: two case studies. Research in the Teaching of English 10, 239-58.

Purves, A. C. (1973). Literature education in ten countries. Stockholm: Almqvist & Wicksell.

Purves, A. C. (1979). That sunny dome: those caves of ice: a model for research in reader response. College English 40, 802-812.

Purves, A. C. (1981). Reading and literature. Urbana, IL: National Council of Teachers of English.

Richards, I. A. (1929). Practical criticism. New York: Harcourt, Brace & World.

Squire, J. R. (1964). The responses of adolescents while reading four short stories. NCTE Research Report #2. Champaign, IL: National Council of Teachers of English.

Squire, J. R., (Ed.) (1968). Response to literature: papers relating to the anglo-american seminar on the teaching of english at dartmouth college, new hampshire 1966. Urbana, IL: National Council of Teachers of English.

Note

The author acknowledges the support of Dean Paul Longo of the Center for the Improvement of Education at Queens College / CUNY and the assistance of Suzann McKiernan and Robin Feingold in developing the instrument and collecting the data.

Table 1. Matrices of Best / Worst Choices: "The Eagle"

Grade 5

Total chi-square (df = 9) = 36.99 $p \leq .001$

Best \ Worst	0	R	M	D	Row Totals	%
0		13	3	3	19	19.59
R	3		6	6	15	15.46
M	2	15		7	24	24.74
D	4	28	7		39	40.21
Column Totals	9	56	16	16	97	Total
%	9.28	57.73	16.49	16.49		

Grade 7

Total chi-square (df = 9) = 45.03 $p \leq .001$

Best \ Worst	0	R	M	D	Totals	%
0		8	4	7	19	20.21
R	5		11	8	24	25.53
M	5	7		7	19	20.21
D	10	19	3		32	34.04
Totals	20	34	18	22	94	Total
%	21.28	36.17	19.15	23.4		

Grade 9

Total chi-square (df = 9) = 38.04 $p \leq .001$

Best \ Worst	0	R	M	D	Totals	%
0		8	3	5	16	18.18
R	2		3	4	9	10.23
M	3	5		2	10	11.36
D	6	42	5		53	60.23
Totals	11	55	11	11	88	Total
%	12.5	62.5	12.5	12.5		

Grade 11

Total chi-square (df = 9) = 33.43 $p \leq .001$

Best \ Worst	0	R	M	D	Totals	%
0		10	7	13	30	36.59
R	4		1	5	10	12.2
M	2	7		4	13	15.85
D	5	17	7		29	35.37
Totals	11	34	15	22	82	Total
%	13.41	41.46	18.29	26.83		

(0 = Original; R = Altered Rhyme; M = Altered Meter; D = Altered Diction)

Table 2. Matrices of Best / Worst Choices: "We Never Know How High"

Grade 5

Total chi-square (df = 9) = 38.79 p ≤ .001

Best \ Worst	0	R	M	D	Row Totals	%
0		11	7	8	26	26.53
R	7		9	4	20	20.41
M	1	6		6	13	13.27
D	10	16	13		39	39.8
Column Totals	18	33	29	18	98	Total
%	18.37	33.67	29.59	18.37		

Grade 7

Total chi-square (df = 9) = 38.2 p ≤ .001

Best \ Worst	0	R	M	D	Totals	%
0		12	4	5	21	22.58
R	6		13	8	27	29.03
M	3	6		5	14	15.05
D	5	10	16		31	33.33
Totals	14	28	33	18	93	Total
%	15.05	30.11	35.48	19.35		

Grade 9

Total chi-square (df = 9) = 43.54 p ≤ .001

Best \ Worst	0	R	M	D	Totals	%
0		13	4	4	21	24.14
R	6		5	4	15	17.24
M	7	1		4	12	13.79
D	6	22	11		39	44.83
Totals	19	36	20	12	87	Total
%	21.84	41.38	22.99	13.79		

Grade 11

Total chi-square (df = 9) = 41.14 p ≤ .001

Best \ Worst	0	R	M	D	Totals	%
0		18	8	4	30	37.5
R	6		3	2	11	13.75
M	0	9		5	14	17.5
D	4	17	4		25	31.25
Totals	10	44	15	11	80	Total
%	12.5	55.0	18.75	13.75		

(0 = Original; R = Altered Rhyme; M = Altered Meter; D = Altered Diction)

Table 3. Matrices of Best / Worst Choices: "The Pasture"

Grade 5

Total chi-square (df = 9) = 49.81 $p \leq .001$

Best \ Worst	0	R	M	D	Row Totals	%
0		14	6	6	26	26.53
R	3		4	3	10	10.2
M	2	19		14	35	35.71
D	5	6	16		27	27.55
Column Totals	10	39	26	23	98	Total
%	10.2	39.8	26.53	23.47		

Grade 7

Total chi-square (df = 9) = 40.55 $p \leq .001$

Best \ Worst	0	R	M	D	Totals	%
0		12	14	9	35	37.63
R	3		7	2	12	12.9
M	3	13		6	22	23.66
D	10	7	7		24	25.81
Totals	16	32	28	17	93	Total
%	17.2	34.41	30.11	18.28		

Grade 9

Total chi-square (df = 9) = 34.08 $p \leq .001$

Best \ Worst	0	R	M	D	Totals	%
0		5	7	11	23	27.06
R	2		2	3	7	8.24
M	5	11		10	26	30.59
D	10	8	11		29	34.12
Totals	17	24	20	24	85	Total
%	20.0	28.24	23.53	28.24		

Grade 11

Total chi-square (df = 9) = 28.9 $p \leq .001$

Best \ Worst	0	R	M	D	Totals	%
0		10	17	3	30	37.5
R	2		4	1	7	8.75
M	3	7		5	15	18.75
D	4	7	17		28	35.0
Totals	9	24	38	9	80	Total
%	11.25	30.0	47.5	11.25		

(O = Original; R = Altered Rhyme; M = Altered Meter; D = Altered Diction)

Table 4. Matrices of Best / Worst Choices: Grade Level Totals

Grade 5

Total chi-square (df = 9) = 109.84 $p \leq .001$

Best \ Worst	0	R	M	D	Row Totals	%
0		38	16	17	71	24.23
R	13		19	13	45	15.36
M	5	40		27	72	24.57
D	19	50	36		105	35.84
Column Totals	37	128	71	57	293	Total
%	12.63	43.69	24.23	19.45		

Grade 7

Total chi-square (df = 9) = 99.19 $p \leq .001$

Best \ Worst	0	R	M	D	Totals	%
0		32	22	21	75	26.79
R	14		31	18	63	22.5
M	11	26		18	55	19.64
D	25	36	26		87	31.07
Totals	50	94	79	57	280	Total
%	17.86	33.57	28.21	20.36		

Grade 9

Total chi-square (df = 9) = 93.64 $p \leq .001$

Best \ Worst	0	R	M	D	Totals	%
0		26	14	20	60	23.08
R	10		10	11	31	11.92
M	15	17		16	48	18.46
D	22	72	27		121	46.54
Totals	47	115	51	47	260	Total
%	18.08	44.23	19.62	18.08		

Grade 11

Total chi-square (df = 9) = 85.82 $p \leq .001$

Best \ Worst	0	R	M	D	Totals	%
0		38	32	20	90	37.19
R	12		8	8	28	11.57
M	5	23		14	42	17.36
D	13	41	28		82	33.88
Totals	30	102	68	42	242	Total
%	12.4	42.15	28.1	17.36		

(0 = Original; R = Altered Rhyme; M = Altered Meter; D = Altered Diction)